Chaofan Lin

siriusneo@sjtu.edu.cn · siriusneo · chaofanlin.com

Education

Tsinghua University, Beijing, China — Sep. 2024 – Unknown

Incoming Ph.D. in Computer Science. Advisor: Mingyu Gao

- A member of Institute for Interdisciplinary Information Sciences (IIIS).

Shanghai Jiao Tong University, Shanghai, China — Aug. 2020 – Present

Bachelor in Computer Science. Advisor: Yong Yu, Weinan Zhang

- A member of **ACM Honors Class**, an elite CS program in SJTU.
- **GPA:** 94 / 100 | **Ranking: 1** / 35.
- **Selected Courses:** Compiler 100/100, Operating System 100/100, Machine Learning 97/100, Mathematical Logic 100/100, Advanced Compiler 100/100, Algorithm 98/100. (And other 20 A+ courses)

My research interest mainly lies in computer systems and architecture, specially in optimizing machine learning and other data intensive applications.

Selected Publications

[1] **Chaofan Lin**, Zhenhua Han, Chengruidong Zhang, Yuqing Yang, Fan Yang, Chen Chen, Lili Qiu. Parrot: Efficient Serving of LLM-based Applications with Semantic Variable. *In 18th USENIX Symposium on Operating Systems Design and Implementation (OSDI 24)*

Research Experience

System and Networking Group, Microsoft Research Asia — 2023.6 - 2024.6

Advised by Zhenhua Han and Yuqing Yang. Topic: LLM Serving System.

- Lead the project Parrot [1], which is a serving system for LLM applications. Parrot proposes Semantic Variable, a unified abstrac- tion to expose application-level knowledge to public LLM services, hence opens a brand new optimization space for LLM inference serving.

Catalyst, Carnegie Mellon University *Research Intern, Remote* — 2022.7 - 2023.5

Advised by Tianqi Chen. Topic: Deep Learning Compiler.

- Developed a training workflow for Relax (the next-gen graph-level IR of TVM), including registration mechanism of operator gradients and an automatic differentiation pass.
- I helped to deploy a simple LLM on device through machine learning compilation (MLC). And also contributed to a project which aims to perform fine-tuning (LoRA) on device.
- Author of 20+ PRs to apache/tvm.

Talks

Cross Platform Training Using Automatic Differentiation on Relax IR — 2023

At TVM Conference 2023. [Video Record]

Honors and Awards

Scholarships

- National Scholarship. (*Top 0.2% national-wide.*) — 2022, 2023
- Foresight-Sequoia Talent Development Fund. (*5 winners each year in SJTU.*) — 2021
- Zhiyuan Undergraduate Excellence Award. *A-level, the highest.* — 2021, 2022
- Zhiyuan Honorary Scholarship. — 2020, 2021, 2022

Competitions

- The Chinese Mathematics Competitions (Shanghai Region). *First Prize.* 2022
- Mathematical Contest In Modeling and Interdisciplinary Contest In Modeling. *Meritorious Winner.* 2021

Selected Projects

ACMClass Online Judge 2022

Previous maintainer of ACMClass OJ, an online judge system for students in SJTU.

Masterball *Course Project of Compiler Design* 2021

A toy compiler implemented in Java, from Mx* (a C++ and Java-like language) to RISC-V assembly, with many optimizations in LLVM IR level, it has a performance close to GCC O2 on testcases. I also implemented a interpreter of LLVM IR with simple Just-In-Time (JIT) technique supported. It received a **perfect score** in two different compilation courses.

NightWizard *Course Project of Computer Architecture* 2021

A RISC-V CPU implemented in Verilog HDL, using Tomasulo algorithm for dynamic scheduling.

CoconutJVM 2022

A toy JVM (Java Virtual Machine) written in C++.

fscape *Course Project of Operating System* 2022

A game based on a simple self-implemented FUSE filesystem. Use 'cd' to move, find a specified file in the file system to escape from this file system maze.

Teaching

Advanced Compiler, Shanghai Jiao Tong University Spring, 2023

Teaching Assistant With Prof. Yong Yu

I gave several talks on Polyhedra model and Register Allocation in this course. [Lecture Notes]

Mathematical Logic, Shanghai Jiao Tong University Fall, 2022

Teaching Assistant With Prof. Qiang Yin and Yijia Chen

Programming Design (A), Shanghai Jiao Tong University Fall, 2021

Teaching Assistant With Prof. Huiyu Weng

I help students in this course implement a simple Python 3 Interpreter.

Skills

- Programming Languages: Python, C/C++, Java, Verilog, Go, Web (HTML, CSS, JavaScript), LaTeX.
- Deep Learning Framework: Pytorch, Tensorflow.
- Some knowledge about FP (Functional Programming), related languages (Haskell, SML) and Type Theory.
- English: CET-6 600, CET-4 661.

Last Updated: May, 2024